#3

0125-0016

Modifications of The VEGF Receptor-2 Protein And Methods Of Use

Inventors: Michele A. McTigue
John A. Wickersham
Chris Pinko
Richard Showalter
Camran V. Parast
Anna Tempcyzk-Russel
Michael R. Gehring
Barabara Mroczkowski
Chen-chen Kan
J. Ernest. Villafranca
Krzysztof Appelt

TECHNICAL FIELD AND INDUSTRIAL APPLICABILITY OF INVENTION

The present invention discloses the isolation of a key portion of the catalytic kinase region of vascular endothelial growth factor receptor 2 or VEGFR-2 through cloning, sequencing and x-ray crystallography. Also disclosed is the deletion of various amino acid residues from an area of the catalytic region called the kinase insert domain (KID). The resulting polypeptide retains comparable *in vitro* kinase activity to that of the wild-type KID and is not necessary for the catalytic activity of the polypeptide, and more importantly, allows complete crystallization of the protein such that it may be characterized by X-ray crystallography. The present invention further discloses x-ray crystallography data useful for identification and construction of therapeutic compounds in the treatment of various disease conditions associated with VEGFR-2.

BACKGROUND OF THE INVENTION

Many physiological events including embryogenesis, organ development, estrus, and wound healing require vascular growth and remodeling (Folkman et al., (1992) J. Biol. Chem. *267*, 10931-10934; Risau, W. (1995) FASEB J. *9*, 926-933.). In addition to these beneficial processes, angiogenesis is also involved in the proliferation of disease states such as tumor growth, metastasis, psoriasis, rheumatoid arthritis, macular degeneration and retinopathy (Pepper, M.S., (1996) Vasc. Med. *1*, 259-266; Kuiper et al., (1998) Pharmacol. Res. *37*, 1-16, 1998; Kumar and Fidler, (1998) In Vivo *18*, 27-34; Szekanecz et al., (1998) J. Investig. Med. *46*, 27-41; Tolentino and Adamis, (1988) Int. Ophthalmol. Clin. *38*, 77-94. Of the signaling pathways known to influence vascular formation, these involving vascular endothelial growth factor (VEGF) haves been shown to be essential and selective for vascular endothelial cells (Dvorak et al., (1995) Am. J. Path. *146*, 1029-1039; Thomas, K., (1996) Cell *271*, 603-606; Ferrara N. and Davis-Smyth, (1997) Endocrine Rev. *18,* 4-25). The therapeutic potential of inhibiting the VEGF pathway has been directly demonstrated by anti-VEGF monoclonal antibodies which were active against a variety of human tumors (Borgström et al, (1996) Cancer Res. *56*, 4032-4039) and ischemic retinal disease (Adamis et al., (1996) Arch. Ophthalmol. *114*, 66-71).

Normal vasculogenesis and angiogenesis play important roles in a variety of physiological processes such as embryonic development, wound healing, organ regeneration and female reproductive processes such as follicle development in the corpus luteum during ovulation and placental growth after pregnancy (Folkman & Shing, 1992). Uncontrolled vasculogenesis and/or angiogenesis has been associated with diseases, such as diabetes, as well as malignant solid tumors that rely on vascularization for growth. Klagsburn & Soker, (1993)

Current Biology 3(10):699-702; Folkham, (1991) J. Natl., Cancer Inst. 82:4-6; Weidner, et al., (1991) New Engl. J. Med. 324:1-5.

Several polypeptides with *in vitro* endothelial cell growth promoting activity have been identified. Examples include acidic and basic fibroblastic growth factor (FGF), vascular endothelial growth factor (VEGF)and placental growth factor. Unlike FGF, VEGF has recently been reported to be an endothelial cell specific mitogen (Ferrara & Henzel, (1989) Biochem. Biophys. Res. Comm. 161:851-858; Vaisman et al., (1990) J. Biol. Chem. 265:19461-19566).

Thus, identification of the specific receptors to which VEGF binds is important to understanding of the regulation of endothelial cell proliferation. Two structurally related tyrosine kinases have been identified to bind VEGF with high affinity: the flt-1 receptor (Shibuya et al., (1990) Oncogene 5:519-524; De Vries et al., (1992) Science 255:989-991) and the KDR/FLK-1receptor, discussed herein. Consequently, it had been surmised that RTKs may have a role in the modulation and regulation of endothelial cell proliferation.

Recent disclosures, such as information set forth in U.S. Patent Application Ser. Nos. 08/193,829, 08/038,596 and 07/975,750,strongly suggest that VEGF is not only responsible for endothelial cell proliferation, but also is the prime regulator of normal and pathological angiogenesis. See generally, Klagsburn & Soker, (1993) Current Biology 3:699-702; Houck, et al., (1992) J. Biol. Chem. 267:26031-26037.

VEGF is a homodimeric cytokine that is expressed in at least four splice-variant forms of 121-206 residues (Ferrara and Davis-Smyth, 1997). Vascular endothelial cells express at least two high-affinity receptors for VEGF: VEGF-R1/Flt-1 and VEGFR-2/KDR. VEGF-R1 and VEGFR-2 are receptor tyrosine kinases each comprised of an extracellular domain that contains 7 immunoglobulin-like segments and binds VEGF, a short membrane spanning region, and a cytosolic domain possessing tyrosine kinase activity. The kinase domain directly follows the extracellular and juxtamembrane regions and is itself followed by another domain (post-kinase domain), which may function in binding of other proteins for signal transduction. These two receptors appear to have different signaling pathways and functions with VEGFR-2 being of primary importance in mitosis of endothelial cells (Waltenberger et al., (1994) J. Biol. Chem. *269*, 26988-26995; Seetharm et al., (1995) Oncogene *10*, 135-147; Shalaby et al., (1995) Nature *376*, 576-579).

Both FGF and VEGF are potent angiogenic factors which induce formation of new capillary blood vessels. Transfection of human breast carcinoma cell line MCF-7 with FGF resulted in cell lines that form progressively growing and metastatic tumors when injected (s.c.) into nude mice. FGF may play a critical role in progression of breast tumors to an estrogen-independent, anti-estrogen resistant metastatic phenotype (McLeskey et al., (1993) Cancer Res. 53: 2168-2177). Breast tumor cells exhibited increased neovascularization, increased spontaneous metastasis and more rapid growth in vivo than did the non-transfected tumors. FGF has been shown to be transforming in NIH-3T3 cells and implicated in tumorigenesis and metastasis of mouse mammary tumors. FGF overexpression conferred a tumorigenic phenotype on a human adrenal carcinoma cell line suggesting that FGF's may also play a role in the transformation of epithelial cells. Polyclonal neutralizing antibodies to FGF inhibited tumor growth in Balb/c nude mice transplanted with K1000 cells (transfected with the leader sequence of bFGF) which form tumors in these mice (Hori et al., (1991) Cancer Res. 51: 6180-9184).

Due to the role of FGF in neovascularization, tumorigenesis and metastasis, there is a need in the art for FGF inhibitors as potent anti-cancer agents that exert their anti-FGF activity by preventing intracellular signaling of FGF.

VEGF, by contrast, is an endothelial cell-specific mitogen and an angiogenesis inducer that is released by a variety of tumor cells and expressed in human tumor cells *in situ*. Unlike FGF, transfection of cell lines with a cDNA sequence encoding VEGF, did not promote transformation, but did facilitate tumor growth *in vivo* (Ferrara, N., and Davis-Smyth, T. (1997)). Furthermore, administration of a polyclonal antibody which neutralized VEGF also inhibited growth of human rhabdomyosarcoma, glioblastoma multiforme and leiomyosarcoma cell lines in nude mice (Kim et al., (1993) Nature 362: 841-843).

In view of the importance of receptor tyrosine kinases (RTKs) to the control, regulation and modulation of endothelial cell proliferation and potentially vasculogenesis and/or angiogenesis, many attempts have been made to identify RTK "inhibitors" using a variety of approaches, including the use of mutant ligands (U.S. Patent No. 4,966,849), soluble receptors and antibodies (Application No. WO 94/10202; Kendall & Thomas, (1994) Proc. Natl. Acad. Sci. 90:10705-09; Kim, et al., 1993), RNA ligands (Jellinek, et al., (1994) Biochemistry 3:10450-56), protein kinase C inhibitors (Schuchter, et al., (1991) Cancer Res. 51:682-687); Takano, et al., (1993) Mol. Bio. Cell 4:358A; Kinsella, et al., (1992) Exp. Cell Res. 199:56-62; Wright, et al., (1992) J. Cellular Phys. 152:448-57) and tyrosine kinase inhibitors (WO 94/03427; WO

92/21660; WO 91/15495; WO 94/14808; U.S. Pat. No. 5,330,992; Mariani, et al., (1994) Proc. Am. Assoc. Cancer Res. 35:2268).

More recently, attempts have been made to identify small molecules which act as tyrosine kinase inhibitors. For example, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642), vinylene-azaindole derivatives (PCT WO 94/14808) and 1-cycloproppyl-4-pyridyl-quinolones (U.S. Pat. No. 5,330,992) have been described generally as tyrosine kinase inhibitors. Styryl compounds (U.S. Pat. No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Pat. No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 AI), selenoindoles and selenides (PCT WO 94/03427), tricyclic polyhydroxylic compounds (PCT WO 92/21660) and benzylphosphonic acid compounds (PCT WO 91/15495) have been described as compounds for use as tyrosine kinase inhibitors for use in the treatment of cancer. None of these compounds, however, have been previously associated with the enzymatic function of the VEGFR-2 receptor. Likewise, none of these compounds have been associated with regulation of vasculogenesis and/or angiogenesis.

Therefore, there is a need in the art to develop small molecule antagonists of the PDGF, FGF, EGF and VEGF pathways individually or as a group. Moreover, if these cytokines signal through a common second messenger pathway within the cell, such antagonists will have broad therapeutic activity to treat or prevent the progression of a broad array of diseases, such as coronary restenosis, tumor-associated angiogenesis, atherosclerosis, autoimmune diseases, acute inflammation, certain kidney diseases associated with proliferation of glomerular or mesangial cells, and ocular diseases associated with retinal vessel proliferation. The present invention was made by discovering a common signaling mechanism, a group of active therapeutic agents, shown to be active by a large number of and variety of predictive assays, and discovering a common intracellular signaling intermediate.

Based on sequence homology and overall domain structure, VEGFRs belong to the platelet-derived growth factor receptor family (PDGFR) which also includes PDGFRα, PDGFRβ, the stem cell growth factor receptor (c-kit), and the colony stimulating factor-1 receptor (CSF-1R/c-fms) (van der Geer et al., (1994) Ann. Rev. Cell Biol. *10*, 251-337). Compared to other protein kinases, members of this family contain an insert of approximately 65-97 residues, termed the kinase insert domain (KID), within the catalytic kinase domain relative to other protein kinases. Within the PDGFR family the KIDs are of varying length and low sequence homology. Deletion or mutation of the KID from PDGFRα, PDGFRβ, c-kit, and CSF-1R have indicated that

this domain is not necessary for intrinsic kinase activity but that it is important for the binding of other proteins involved in signal transduction, via autophosphorylation of KID tyrosine residues (Taylor et al., (1989) EMBO J. *8*, 2029-2037; Heidaran et al., (1991) Mol. Cell. Biol. *11*, 134-142; Yu et al., (1991) Mol. Cell. Biol. *11*, 3780-3785; Kazlauskas et al., (1992) Mol. Cell. Biol. *12*, 2534-2544; Lev et al., (1992) Proc. Natl. Acad. Sci. USA *89*, 678-682; Reedjik et al., (1992) EMBO J. *11*, 1365-1372; Bazenet et al., (1996) Mol. Cell. Biol. *16*, 6926-6936). Although the signaling pathways and the specific role of the KID are still not fully determined for VEGFRs, the VEGFR-2 KID does contain two tyrosines which are known to be autophosphorylation sites (Dougher-Vermazen et al., (1994) Biochem. Biophys. Res. Comm. *205*, 728-738).

Since the determination of the first cyclic AMP-dependent protein kinase (cAPK) structure (Knighton et al., (1991) Science *253*, 407-413) a variety of protein kinase structures have been reported (reviewed in Johnson et al., (1996) Cell *85*, 149-158). Among the receptor protein tyrosine kinases (RTKs), structures of the kinase domain of the insulin receptor (IRK) (Hubbard, et al., (1994) Nature *372*, 746-754; Hubbard, (1997) EMBO J. *16,* 5572-5581) and the fibroblast growth factor receptor-1 (FGFR1) (Mohammadi et al., (1996) Cell *86*, 577-87; Mohammadi et al., (1997) Science *276*, 955-960) have been determined.

SUMMARY OF THE INVENTION

The present invention discloses the generation, kinetic characterization, and structure determination of a modified kinase domain of the VEGFR-2 protein, containing 18 residues of the 68 residue KID. This 2.4 Å crystal structure of the phosphorylated VEGFR-2 catalytic domain is the first reported structure of a kinase domain of the PDGFR family. This structure provides insights into the orientation of the KID domain of VEGFR-2 which may be relevant to other PDGFR family members. Furthermore, as inhibition of VEGFR-2 kinase has broad clinical applications, this structure provides a three-dimensional description of the target for structure-based design of small molecule VEGFR-2 inhibitors as therapeutic agents.

It is an object of the present invention to disclose an effective method for screening candidate compounds that are specifically agonists or antagonists of various proteins which can be included in the receptor tyrosine kinase family (RTK) by crystallizing RTKs and particularly the VEGFR-2 receptor in order to use molecular modeling of the x-ray crystallography data to model the binding of candidate compounds.

There is disclosed a method for designing and screening potentially therapeutic compounds with activities such as: (1) inhibiting new blood vessel formation that is useful for treating or preventing progression of diabetic retinopathy, cavernous hemangiomas, Kaposi's sarcoma, tumors composed of endothelial-like cells, and growth of cancer cells by preventing their development of a new blood supply: (2) suppressing development of kidney diseases due to cytokine induced proliferation of mesangial cells and/or glomerular epithelial cells that is useful for treating or preventing progression of diabetic glomerulosclerosis and other glomerulonephritis of various types and etiologies; (3) preventing joint destruction accompanying rheumatoid arthritis due to proliferation of synovial cells; (4) suppressing manifestations of psoriasis due to proliferation of keratinocytes and accumulation of inflammatory cells; (5) suppressing accelerated atherogenesis involved in restenosis of coronary vessels or other arterial vessels following angioplasty; (6) suppressing atherogenesis, coronary artery disease and other vasculopathies due to atherogenesis; and (7) suppressing tumor growth via paracrine or autocrine mediated responses to other cytokines such as PDGF, FGF EGF or VEGF that is useful for treating or preventing progression of tumors such as breast cancer stimulated through overexpression of her-2-neu receptor, wherein the inventive method comprises administering a compound that inhibits signal transduction.

The present invention is useful in developing methods that are used in the iterative drug design process. The process identifies potential agonists and antagonists to VEGFR-2 by *de novo* design of novel drug candidate molecules which bind to the VEGFR-2 receptor to improve their potency. The x-ray crystallographic coordinates disclosed herein, will allow generation of 3-dimensional models of the catalytic site and drug binding site of the VEGFR-2 protein.

De novo design primarily consists of the generation of molecules via the use of computer programs which build and link fragments or atoms into a site based upon steric and electrostatic complementarity, without reference to substrate analog structures. The drug design process begins after the structure of a target RTK is solved to at least a resolution of 2.8 Å. Refinement of the structure to a resolution of 2.5 Å or better, with "fixed" water molecules in place provides more optimal conditions to undertake drug design.

It is another object of this invention to identify KIDs of proteins in the RTK family and develop deletions in said KIDs such that the proteins will be crystallizable and suitable for measurement by x-ray crystallographic means.

It is a further object of this invention to disclose a process whereby KID regions from a member of the RTK family of genes such as PDGF, EGF, VEGF and others are modified by deletion of amino acids from the KID regions so as to impart favorable physical characteristics of the resulting polypeptide product. Examples of such favorable physical characteristics are increased solubility, greater stability to temperature variations making the polypeptide suitable for analysis by nuclear magnetic resonance, high throughput screening, biochemical characterizations, x-ray crystallography, calorimetry and other diagnostic means.

It is yet another object of this invention to developing screening methods used in the drug design process of potential agonists and antagonists to proteins in the RTK family by de novo design of novel drug candidate molecules with potentially nanomolar potencies. The x-ray crystallographic coordinates disclosed based on the deletion mutated KIDs and various other deletions of said proteins in the RTK family, will allow generationof 3-dimensional models of the active binding site of the proteins in the RTK family.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Secondary structure assignments (as given by Procheck) for the catalytic domain of VEGFR2 and sequence alignment with other representative receptor tyrosine kinases. α helices are designated as αB-αI, β strands are designated as β1-β8. The site of 50 residue deletion in VEGFR2Δ50 is indicated by |. The site of the E990V mutation in VEGFR2D50 is denoted by an *. Sequences are from: VEGFR2 (reported here); FGFR1 (Swiss protein database #P11362); IRK (EMBL protein database #A18657; numbering as in Mohammadi et al., 1996); VEGFR1 (Swiss protein database #P17948); PDGFRα (Swiss protein database #P17948).

Figure 2. Overall fold of VEGFR2Δ50P, FGFR1, and IRKP.
Backbone representation of structures of the kinase domains of (A) VEGFR2 (VEGFR2Δ50P), (B) FGFR1 (molecule A of PDB entry 1FGK, Mohammadi et al., 1996), and (C) IRKP (PDB entry 1IR3, Hubbard et al., 1997). The views shown in A, B, and C are identical views generated from superpositions of the C-terminal domains. The positions of the termini are denoted by N and C. The nucleotide-binding loop (orange), kinase insert domain (pink), and activation loop (yellow) are highlighted. In (C) the bound AMP-PNP is shown in green and the peptide substrate is shown in red. Figure prepared with INSIGHT II.

Figure 3. Catalytic site of VEGFR2Δ50P and IRKP.

Cross section of the catalytic site of (A) VEGFR2Δ50P and (B) IRKP (PDB entry 1IR3; Hubbard et al., 1997) structures. Atoms are colored by element type: carbon (green), oxygen (red), nitrogen (blue), sulfur (yellow), phosphorous (pink),and magnesium ion (orange). (A) includes only protein atoms. (B) includes protein atoms, AMP-PNP atoms, and Mg^{2+} ions. Figure generated using INSIGHT II.

Figure 4. Nucleotide binding site of VEGFR2Δ50P and FGFR1.
Stereo view showing Cα trace and some sidechains of a superposition of the nucleotide binding sites of the VEGFR2Δ50P and the FGFR1-(AMP-PCP) complex (molecule B, Mohammadi et al., 1996) structures. The superposition was done using Cα positions of helices (D,E,F,G,H, and I) of the C-terminal lobes. Carbon atoms of VEGFR2Δ50P are shown in yellow and carbon atoms of FGFR1 are shown in purple. The coloring for other protein atoms is: oxygen (red), nitrogen (blue), and sulfur (green). The AMP-PCP in the FGFR1 structure is depicted in orange. Labels correspond to VEGFR2Δ50P residues. Figure created with Xfit (McRee et al., (1992) J. Mol. Graph. *10*, 44-46.).

Figure 5. Electron density map of the kinase insert domain area of VEGFR2Δ50P.
Stereo view of a $2F_O$-F_C map computed at 2.4Å and contoured at 1.2δ and superimposed with the refined model. Carbon atoms are yellow, oxygen atoms red, and nitrogen atoms are blue. Water molecules are depicted as red crosses. Figure created with Xfit (McRee et al., 1992).

Figure 6. Kinase insert domain of VEGFR2Δ50P. Stereo cross section showing the ordered residues of the kinase insert domain of VEGFR2Δ50P. Carbon atoms are yellow, oxygen atoms are red, nitrogen atoms are blue, and sulfur atoms are green. View is rotated roughly 180° from Figure 5. Figure created with Xfit (McRee et al., 1992).

Figure 7. Resulting X-ray crystallography coordinates for VEGFR-2 based on the method disclosed in the crystallization and data collection section.

DETAILED DESCRIPTION AND PREFERRED EMBODIMENTS OF THE INVENTION

Cloning of The VEGFR-2 Protein

The coding sequence (Terman et al., (1992) Biochem Biophys. Res. Commun. *187*, 1579-86) for the cytoplasmic domain of the VEGFR-2 was amplified by PCR (Mullis et al., (1986). Biotechnology *24*, 17-27) from a human aorta cDNA pool (Clontech Palo Alto, CA). Two

overlapping sequences were amplified independently. Vcyt (residues M806-V1356), which represented the entire cytoplasmic domain, and Vcat (residues C817M-G1191), with boundaries based upon a primary amino acid sequence alignment with the insulin receptor kinase catalytic domain (Wei et al., (1995) J. Biol. Chem. *270*, 8122-8130).

The PCR oligonucleotide primer sequences for Vcyt were:
Vcyt5 5'-CAGCATATGGATCCAGATGAACTCCCATTGG3' (Seq. ID No. 1) and
Vcyt3 5'-GCGGTCGACTTAAACAGGAGGAGAGCTCAGTGTG3' (Seq. ID No. 2).

The PCR oligonucleotide primer sequences for the Vcat were:
Vcat5 5'-GCACATATGGAACGACTGCCTTATGATGCCAGG3' (Seq. ID No. 3) and
Vcat3 5'-CCTGTCGACTTATCCAGAATCCTCTTCCATGCTCAAAG5' (Seq. ID No. 4).

The amplified DNA was digested with the restriction enzymes NdeI and SalI, ligated into the *E coli* plasmid pET24a (Novagen Madison, WI) and sequence verified. When compared to the original VEGFR-2 sequence in Genbank, (Accession number 346345) two nucleotide differences were noted that resulted in codon changes (Glu848-Val and Asn835-Lys) in both Vcyt and Vcat. Our sequence agrees with subsequent VEGFR-2 Genbank submissions (Accession numbers 2655412 and 3132833).

Mutations were introduced by oligonucleotide site directed mutagenesis (Kunkel, 1985) using the Muta-Gene in vitro Mutagenesis Kit from (Bio-Rad Hercules, CA). The Vcat DNA fragment was subcloned from the pET24a vector using an NdeI-XhoI digest into the vector pMGH4 (Schoner et al., 1986, Kan et al., 1992) and this vector was used to generate the ssDNA uracil template (minus strand) in *E. coli* strain CJ236 supplied in the kit. An oligo (5'-CTCAGCAGGATTGATAAGACTACATTGTTG3') was designed to create a construct (Vcat(ΔG1172-G1191)) which truncated the C-terminus to residue D1171. Another oligo (5'-GAATTTGTCCCCTACAAGGAAGCTCCTGAAGATCTG3') was designed to delete the central 50 residues (residues T940-E989) of the insert kinase domain, based on a sequence alignment with FGFR1 (Mohammadi et al, 1996). Sequence analysis detected an inadvertent Glu990-Val mutation. All DNA modification and restriction enzymes were purchased from New England Biolabs and oligonucleotides were purchased form Genosys Biotechnology.

The VEGFR2Δ50 construct was made in several steps to combine the necessary mutations into the baculovirus expression vector pAcSG2 (Pharmingen San Diego, CA). Step 1; the coding region for Vcyt was PCR subcloned from the pET24a vector into the NcoI-KpnI sites of vector pAcSG2. Step2; a

2358bp ScaI-BglII DNA fragment from plasmid pMGH4-Vcat (ΔT940-E989,E990V) was ligated to a 1695bp BglII-ScaI DNA fragment from pMGH4-Vcat (ΔG1172-G1191) creating a pMGH4-Vcat (ΔT940-E989,E990V,ΔG1172-G1191) vector. Step 3; a 913bp BstEII-EagI DNA fragment a pMGH4-Vcat (ΔT940-E989,E990V,ΔG1172-G1191) was ligated to a 3290bp EagI-BstEII DNA fragment from pAcSG2-Vcyt creating pAcSG2-Vcyt (ΔT940-E989,E990V,ΔG1172-G1191), also referred to as VEGFR2Δ50. This final construct was sequenced verified through the entire coding region and confirmed to contain only these known mutations from the wild-type sequence (sequence shown in Figure 1).

DNA encoding VEGFR2Δ50 was transfected into Sf9 cells with linearized baculovirus DNA according to the protocol of the manufacturer (Pharmingen San Diego, CA). Single plaques were isolated from this transfection and high titer stocks generated. All stocks were examined by isolation of baculoviral DNA and PCR amplification of the insert using the polyhedron forward and reverse primers (Invitrogen). Sf21 cells were infected at 1-1.5 million cells/mL at MOI=5 for 72 hours and harvested by centrifugation.

Purification Of VEGFR2Δ50 From Sf21 Cells

Cell pellets were lysed by dounce homogenization and sonication in 20 mM Tris pH 8.0, 20 mM NaCl, 5 mM DTT, and 5% (v/v) glycerol. The lysate was centrifuged for 50 minutes at 35,000 rpm in a Ti45 rotor. The soluble fraction was loaded onto a 40 ml Q-30 anion exchange column (Pharmacia) and eluted with a 20 mM to 600 mM NaCl gradient in 20 mM Tris pH 8.0, 5 mM DTT, and 5% (v/v) glycerol over 20 column volumes. VEGFR2Δ50 protein was pooled by SDS-PAGE gel analysis and by the presence of kinase activity as measured against gastrin substrate peptide substrate (Boehringer Mannheim). Pooled material was loaded onto a 40 mL hydroxyapatite (Bio-Rad) column and washed extensively with 20 mM Tris pH 8.0, 50 mM NaCl, 5 mM DTT, and 5% glycerol. Protein was eluted using a 500 mL linear gradient from 0 to 50 mM potassium phosphate pH 8.0, 50 mM NaCl, 5 mM DTT, and 5% glycerol. VEGFR2Δ50 protein was pooled by SDS-PAGE gel analysis and by the presence of kinase activity as measured against the gastrin peptide. Material from this column was then diluted 1:1 with 20 mM Tris pH 8.0, 20 mM NaCl, 5 mM DTT, and 5% glycerol and loaded onto an 8 mL Q-15 anion exchange column (Pharmacia). Protein was eluted using with a 180 mL linear NaCl gradient (20 mM-175 mM) in 20 mM Tris pH 8.0, 5 mM DTT, and 5% glycerol. VEGFR2Δ50 protein was pooled as described above. 4M $(NH_4)_2SO4$ was added to the pool to final concentration of 0.6 M and the pool loaded onto a 10 mL HP-phenyl sepharose column (Pharmacia). VEGFR2Δ50 protein was eluted using a 200 mL linear reverse gradient from 0.6 M to 0 M $(NH_4)_2SO4$ in 20 mM Tris and 5 mM DTT. Purified VEGFR2Δ50 protein was buffer exchanged into 50 mM Hepes pH 7.5, 10

mM DTT, 10 % glycerol, and 25 mM NaCl over a 500 ml G-25 column (Pharmacia) and concentrated to 1 mg protein/mL through a 10 kD cutoff polysulfone membrane (Amicon). Final material was aliquoted and flash frozen in liquid N_2 and stored at -70'C.

Kinetic Assays

The coupled spectrophotometric assays were done with purified VEGFR2Δ50 protein that was autophosphorylated under conditions: protein (4 mM), ATP (3 mM), $MgCl_2$ (40 mM), DTT (5 mM), in Hepes (100 mM), 10% glycerol, pH 7.5 at 4 °C for 1 hour.

Coupled Spectrophotometric Assay for the Forward Direction

Tyrosine kinase assays were monitored using a Beckman DU 650 Spectrophotometer. Production of ADP was coupled to oxidation of NADH using phosphoenolpyruvate (PEP) through the actions of pyruvate kinase (PK) and lactic dehydrogenase (LDH). The oxidation of NADH was monitored by following the decrease in absorbance at 340 nm (e_{340}=6.22 cm^{-1} mM^{-1}). Typical reaction solutions contained: 1 mM PEP, 250 mM NADH, 50 units of LDH/mL, 20 units of PK/mL, 5 mM DTT, in 200 mM Hepes, pH 7.5 and varying concentrations of poly(E_4Y_1) (Sigma), ATP and $MgCl_2$. Assays were initiated with 40 nM of VEGFR2Δ50 protein.

Coupled Spectrophotometric Assay for the Reverse Reaction

ATP generation was coupled to production of NADH via the action of hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6PD). In this assay, HK catalyzes the conversion of ATP to ADP and glucose-6-phosphate. Glucose-6-phosphate is then oxidized to D-6-phosphogluconopyranose-1,5-lactone by G6PD with concomitant reduction of NAD to NADH which can be monitored at 340 nm. Typical assay solution contained: glucose (10 mM), NAD (40 mM), DTT (5 mM), $MgCl_2$ (4 mM), HK (15 unit/mL), G6PD (15 units/mL) and indicated concentrations of ADP and phospho-poly(E_4Y). The reactions were initiated with addition of VEGFR2Δ50 protein (600-900 nM).

Evaluation of Potential Agonists and Antagonists of the VEGFR2Δ50 Protein

Based on the above spectrophotometric and kinetic assays, one can evaluate potential candidate agonists or antagonists of the VEGFR2Δ50 protein by addition of the candidate

compounds to the above assay in a competition. As stated above, the kinetics of the activity of the VEGFR2Δ50 protein were measured against the gastrin peptide. The activity in the presence and absence of a candidate compound is measured and the resulting kinetic data is compared. The affinity of the candidate for the receptor will be reflected in the shift to the right of the kinetic curves indicating a competitive antagonist or with a decrease in the maximum activity, which would indicate a non-competitive antagonism. Conversely, a shift to the left of the kinetic curves would indicate a competitive agonist to the VEGFR2Δ50 protein. See generally, Bourne, H.R., et al. in, (1987) Basic & Clinical Pharmacology (Katzung, et al., eds) (Ch. 3) 9-22.

In Vitro Autophosphorylation Of VEGFR2Δ50 For Crystallization And Mass Spectrometry.

Aliquots of frozen VEGFR2Δ50 protein were thawed by immersion in cold H_20 and pooled at 4'C. $MgCl_2$ and ATP were added to 26 mM and 4 mM, respectively. VEGFR2Δ50 was incubated at 4'C for 1 hour. This material (VEGFR2Δ50P) was then buffer exchanged into a solution of 10 mM Hepes 7.5, 10 mM DTT, and 10 mM NaCl and concentrated using a Centriprep-10 (Amicon) to 5 mg protein/mL.

Mass spectrometry

Trypsin digestion:. Trypsin digestions of purified VEGFR2Δ50 and VEGFR2Δ50P were conducted at 37°C suing 0.37 mg/ML protein in 25 mM NH_4HCO_3 at pH. 7.7 with a reaction volume of 100 µL for two days.

MALDI/MS. MALDI-MS analyses were performed in a Voyager-Elite, time-of-flight mass spectrometer with delayed extraction (PerSeptive Biosystems, Inc., Framingham, MA). A volume of 1 µL of digested protein sample was mixed with 1 µL of matrix (a-cyano-4-hydroxy-cinnamic acid) in a solution of 50% (v/v) solution of acetonitrile and 0.25% (w/w) trifluoroacetic acid in water. Samples were irradiated with a nitrogen laser operated at 337 nm.

NanoESI-MS. NanoESI-MS analyses were performed on a triple quadrapole mass spectrometer (PE Sciex API III, Alberta, Canada) modified with a NanoESI source from Protana A/S, (Denmark). The ESI voltage was set at 700 V and the orifice settings were maintained at 100 V. 3 µL of digested protein was mixed with 7 µL of methanol and 0.5 µL formic acid and then 4 µL of this sample was injected into the mass spectrometer. Ion scans were used to obtain the sequence of phospho-peptides.

Crystallization and Data Collection

Purified phosphorylated VEGFR2Δ50 was concentrated on average to 5 mg protein/mL using a Centricon-10 centrifugal concentrator. Crystals were grown by the hanging drop vapor diffusion method at 4°C. Drops containing 2 µL of protein solution and 2 µL of a mother liquor solution (100 mM Hepes at pH 7.2, 2 M $(NH_4)_2SO_4$,, and 2% (v/v) monomethylether polyethylene glycol mW=550) were equilibrated above a 1 mL reservoir of the mother liquor solution to which 50 mM β-mercaptoethanol had been added. Crystals appeared after 3-4 days and grew to as large as 0.3 x 0.2 x 0.5 mm over 21 days.

X-ray diffraction data sets were collected using a Rigaku RU-200 rotating anode X-ray generator (CuKα) operated at 50 kV and 100 mA and equipped with Supper focusing mirrors and a MAR345 MAR Research image plate detector. Data collection on frozen crystals was done by transferring a crystal into a cryoprotectant solution (100 mM Hepes at pH 7.2, 2.2 M $(NH_4)_2SO_4$, 0.6 M sucrose, 0.55 M glucose, and 2% (v/v) monomethylether polyethylene glycol MW=550), flash freezing the crystal in liquid nitrogen, and then transferring the frozen crystal into a stream of nitrogen at -186 °C. Data was integrated and scaled using DENZO and SCALEPACK (Otwinowski, 1993) Data collection statistics aregiven in Table 2.

Initial protein phases were obtained using the AMoRe molecular replacement program (Navaza, 1994), molecule 1 of the FGFR1 structure (Mohammadi et al., 1996; PDB entry 1FGK) as a search probe, and the native1 data set. The correct solution was achieved by including the FGFR1 sidechains and removing from mobile residues of the activation loop (640-660), the N-terminus (464-467), a short loop (517-520), and the C-terminus (760-762) from the search model. The correct solution was the top peak in the rotation and translation functions with a correlation coefficient of 0.31. Rigid body refinement in AMoRe improved the solution to a correlation coefficient of 0.49 and an *R*-factor of 46.3% in the 12.0 - 4.0 Å resolution range. The correctness of this solution was cross-checked by calculation of a difference Fourier with a $KAu(CN)_2$ derivative. This derivative was generated by soaking a crystal for 3 days in reservoir solution containing 0.5 mM $KAu(CN)_2$ and then increasing the heavy atom concentration to 5 mM and soaking for an additional 64 hours. Scaling of data sets, Patterson calculations, Fourier calculations, and the generation of phases were done using Xtalview (McRee et al., 1992)

Refinement of the model was done using Xplor version 3.1 (Brünger, 1992). Calculation of electron density maps and model fitting was done using XtalView (McRee et al., 1992) Refinement was begun using a data set collected at 4 °C (native2) and was completed using a

data set (native3) collected at -186 °C. The final R-factor is 20.2% for data in the range 8-2.4 Å (Fo > 2δ). The average B value for all atoms is 31.8 $Å^2$ for protein atoms and 42.8 $Å^2$ for water molecules. The final model includes residues 820-939, 998-1047, and 1064-1168; of these residues the sidechains of K838, R842, F845, K939, D998, K1023, R1027, Y1038, K1039, K1110, and E1113 could not be modeled beyond Cα due to a lack of interpretable density. Analysis of main-chain torsion angles as done using PROCHECK (Laskowski et al., 1993) shows of the 275 residues in the model none occur in the disallowed region and only 4 occur in the generously allowed region of a Ramachandran plot. 182 water molecules were fit to electron density peaks which were greater than 3δ and were located in positions to make reasonable hydrogen bonds to the protein or other water molecules.

Superpositions of various kinase structures was done using the graphics program Insight II (Molecular Simulations Inc, San Diego, CA).

Example 1- Structure Determinations

The tyrosine kinase domain of human VEGFR-2 lacking the 50 central residues of the 68 residues of the KID was expressed in a baculovirus/insect cell system. Of the 1356 residues of full-length VEGFR-2 this construct (VEGFR2Δ50) contains residues 806-939 and 990-1171 of the cytosolic domain (Figure 1). VEGFR2Δ50 also contains one point mutation (E990V) within the KID relative to wild-type VEGFR-2.

In addition to catalyzing its autophosphorylation, VEGFR2Δ50 is also capable of catalyzing phosphorylation of a poly(E_4Y) exogenous substrate. Detailed kinetic analysis (Table 1) revealed that its kinetic parameters were nearly identical to that of a comparable VEGFR-2 protein construct containing the entire KID (Parast et al., in press). These results taken together indicate that VEGFR2Δ50 is a fully active functional enzyme. Therefore, deletion of 50 central residues of the KID has no observed effect on the catalytic steps of the phosphotransfer reaction. It was also determined that deletion of more than 60 amino acids from the KID region did cause a diminishment in the activity of the enzyme.

Tabl 1: Kinetic constants of VEGFR2Δ50

	Forward Reaction		
Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}M^{-1}$)
MgATP	0.153	13.3	87×10^3
poly(E_4Y)	2.1		63×10^2
Mg^{2+}	6.8		20×10^2

	Reverse Reaction		
Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}M^{-1}$)
MgADP	0.056	0.13	23×10^2
P-poly(E_4Y)	1.0		13×10^1

The VEGFR-2 KID sequence is hydrophilic and highly charged, containing 6 lysine, 5 arginine, 8 glutamic acid, and 5 aspartic acid residues (Figure 1). Initially several protein constructs containing the VEGFR-2 catalytic domain with the entire KID were generated. After exhaustive attempts to crystallize these protein constructs failed to yield even marginal crystals, the VEGFR2Δ50 construct was created to test the idea that the highly charged KID was interfering with crystallization. As determined by dynamic light scattering this VEGFR2Δ50 construct, which eliminated 14 charged residues, exhibited markedly better stability to temperature and protein concentration than protein constructs containing the entire KID.

For crystallization, purified VEGFR2Δ50 was autophosphorylated in vitro by incubation with MgATP. Matrix-assisted laser desorption ionization (MALDI) and nanoelectrospray ionization (NanoESI) mass spectrometry analysis of full-length phosphorylated VEGFR2Δ50 (VEGFR2Δ50P) protein and tryptically digested peptides indicates phosphorylation of Y1059 using the autophosphorylation conditions described here. Crystals diffracting to 2.2 Å were obtained of VEGFR2Δ50P in an unligated state. The crystals belong to the orthorhombic space

group $P2_12_12_1$ with one VEGFR2Δ50P molecule in the asymmetric unit. Initial crystallographic phases were determined by molecular replacement using the structure of the unphosphorylated kinase domain of FGFR1 (Mohammadi et al., 1996) as a search model. The correctness of the molecular replacement solution was cross-checked using a gold cyanide derivative. The derivative data, however, was not used for phase calculations of electron density maps used to build the model. The structure has been refined to an *R*-factor of 20.2% for 8-2.4Å data (Fo > 2δ). VEGFR2Δ50P residues for which backbone atoms were not modeled due to disorder include the N-terminal residues 806-819, residues 990-997 of the KID, residues 1048-1063 of the activation loop, and residues 1169-1171 of the C-terminus. Structure determination statistics are included in Table 2.

Table 2: VEGFR2Δ50P structure determination statistics

Data Set	Native (3)	Native (1)	Native (2)	$KAu(CN)_2$
Data resolution (Å)	15-2.2	20-3.0	15-2.4	15.3.1
R_{sym} (%)	5.2[a] (19.6)[b]	8.4 (19.2)	7.0 (21.9)	7.1 (19.5)
Completeness (%)	93.0 (81.0)	97.5 (98.4)	98.8 (98.8)	96.5 (95.0)
Temperature (^oC)	-186	room (~21)	4	4
Unit cell a (Å)	95.41	97.10	98.52	97.71
Unit cell b (Å)	96.04	96.94	96.50	96.97
Unit cell c (Å)	38.22	38.63	38.56	38.52
Refinement resolution (Å)	8-2.4	--	--	--
Refined *R* (%)	20.2[c,d]	--	--	--

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$

[b] Value in parenthesis is for highest (resolution shell)

[c] $R = \Sigma_{hkl}||F_O(hkl)| - |F_C(hkl)|| / \Sigma_{hkl}|F_O(hkl)|$

where F_O and F_C are the observed and calculated structure factors, respectively ($F_O > 2\delta$)

[d] Model includes 275 protein residues and 182 water molecules

Overall Kinas Fold

Analogous to previously reported structures of both serine/threonine and tyrosine protein kinases, VEGFR2Δ50P is folded into two lobes with catalysis of phosphotransfer taking place in a cleft between the two lobes (reviewed in Cox et al., 1994; Johnson et al., 1996) A C α trace of the VEGFR2Δ50P structure is shown in Figure 2a. Kinase secondary structural elements are designated (Figure 1) according to the convention originally given for cAPK (Knighton et al., 1991). The N-terminal lobe (approximately residues 820-920) folds into a twisted β sheet with one α helix (αC). The β structure comprises five antiparallel strands (β1-β5), three of which (β1-β3) are highly curved and curl over the other two strands (β4-β5). The larger C-terminal domain (approximately residues 921-313) contains two antiparallel β strands (β7-β8), which lie at the top of the C-terminal domain adjacent to the N-terminal β-sheet. Seven α-helices (αD, αE, αE-F, αG, αH, αI) form the remaining core of the C-terminal domain. Like other kinases, VEGFR2Δ50P contains two functionally important loop regions: the glycine-rich nucleotide binding loop (residues 841-846), the catalytic loop (residues 1026-1033) and the activation loop (residues 1046-1075) (Figures 1 and 2a).

Of the reported kinase structures, the VEGFR2Δ50P structure resembles most closely that of the catalytic domain of FGFR1 (Mohammadi et al., 1996; PDB entry 1FGK) with which it shares approximately 55% sequence identity (Figure 1). Since the two molecules in the crystallographic asymmetric unit of the FGFR1 structure solution are very similar, comparisons to VEGFR2Δ50P will primarily be described only for FGFR1 molecule A. Least squares superposition of 82 Cα positions of (β1-β5) of the N-terminal lobe or 152 Cα positions residues (αD, αE, αF, αG, αH, αI) of the C-terminal lobe between FGFR1 and VEGFR2Δ50P result in respective rms deviations of 0.40 Å and 0.52 Å. A relative rotation of approximately 5° between the two lobes results in the interlobe cleft of VEGFR2Δ50P being slightly larger and more open. Measurement of distances between equivalent Cα's (K523 and R675 of FGFR1, S877 and R1080 of VEGFR2Δ50P) at the ends of the cleft reveal that this distance is 25.3 Å in VEGFR2Δ50P as compared to 23.2 Å in FGFR1. This is however a minor difference, as compared to much larger relative lobe rotations observed among kinase structures in various ligation and phosphorylation states (Johnson et al., (1996) Cell *85*, 149-158). For example, the inter-lobe orientation seen here for VEGFR2Δ50P is in an approximately 20° more open conformation than that seen in the ternary complex structure of the phosphorylated kinase domain of IRK bound to the ATP analog AMP-PNP and a peptide substrate (Hubbard, (1997) EMBO J. *16*, 5572-5581; PDB entry 1IR3) (Figure2c).

While the β-strand positions of the N-terminal lobe agree well between VEGFR2Δ50P and FGFR1, the structures do diverge significantly at the N-terminal residues preceding the first

conserved region starting at residue W827 (Figure 2a and 2b). The first 14 residues (M806-E819) of VEGFR2Δ50P are completely disordered and the next seven residues (L820-R826) form an extended loop structure. It is likely that residues 806-819 do not form part of the active kinase region but are instead part of, or are adjacent to, the juxtamembrane region of VEGFR-2. Residues 820-826 do seem to be part of the kinase domain, although a flexible one, as analogous residues are also ordered in the structures of FGFR1, IRK, and the non-receptor tyrosine kinase Lck (Yamaguchi and Hendrickson, (1996) Nature *384*, 484-489). Other differences between the VEGFR2Δ50P structure and other kinase structures occur at the kinase insert domain and the activation loop (discussed below).

Catalytic loop and ATP binding site

In protein kinases, the loop between αE and β7 has been termed the catalytic loop as it contains an invariant aspartic acid (D1028) that is believed to function as a catalytic base in the phosphotransfer reaction (Johnson et al., 1996). This aspartic acid is part of a stretch of residues (H1026-N1033) whose sequence HRDLAARN is highly conserved among protein tyrosine kinases. In VEGFR2Δ50P the backbone position and most sidechain positions of this loop are similar to those in the unliganded FGFR1 and ternary phosphorylated IRK (IRKP) complex structures. As seen in these previous structures the sidechain carboxylate of the catalytic loop aspartic acid (D1028) is hydrogen bonded to the sidechains of the conserved arginine (R1032) and asparagine (N1033) (Figure 3).

The ATP binding site of protein kinases lies at the cleft between the N and C-terminal lobes (Figure 2c). For VEGFR2Δ50P, the residues forming this site consist primarily of residues E917-N923, joining the two lobes, and residues L840-I849 which include parts of β1, β2, and the glycine-rich loop of G841-G846. The glycine-rich loop, also referred to as the nucleotide binding loop, is a flexible segment whose position differs among kinase structures in various activated and liganded states. In VEGFR2Δ50P this loop is fairly well ordered and all atoms could be modeled with the exception of the sidechains of R842 and F845. The relative position and conformation of this loop is similar to that observed in the unligated FGFR1 structure. However, this position is markedly different from that in the IRKP ternary complex structure in which the approximately 20° relative rotation of the N and C-terminal lobes results in the glycine-rich loop being 5 Å closer to the C-terminal lobe than in VEGFR2Δ50P structure.

In reported kinase structures with bound ATP or an ATP analog, the adenine ring makes two conserved hydrogen bonds with the protein backbone. In the structure of FGFR1 with AMP-

PCP bound (Mohammadi et al., 1996) these hydrogen bonds are between the adenine NH_2 and the backbone C=0 of E562 (E917 VEGFR2Δ50P) and between the adenine N1 and the backbone NH of A546 (C919 VEGFR2Δ50P). Although the structure presented here does not contain a bound nucleotide, the similarities in the positions of these backbone atoms to those in FGFR1 indicate that these hydrogen bonds would be formed in a VEGFR2Δ50P-ATP complex and therefore the adenine is expected to bind in a similar position (Figure 4).

Variation in the ATP-binding sites of kinases involved in disease is of considerable importance in the design of selective ATP-competitive inhibitors as therapeutics. A comparison of the ATP binding sites of FGFR1 and VEGFR2Δ50P reveals that while the overall architecture of the site is conserved, several sequence differences result in differences in the shape of the accessible area for ligand binding. Specific sequence differences between FGFR1 and VEGFR-2 in this site include: V899 (I545 FGFR1), F918 (Y563 FGFR1), C919 (A564 FGFR1), and C1045 (A640 FGFR1) (Figure 4). Similarly, comparison to the ternary IRKP complex structure reveals variation in the adenine site at V916 (M1076 IRK), F918(L1078), C919 (M1079 IRK), L1035 (M1139 IRK), and C1045 (G1149 IRK). Even greater sequence and structural variation in the adenine site is seen when the VEGFR2Δ50P structure is compared to serine/threonine kinase structures, suggesting that these differences are useful in the design of selective ATP-competitive inhibitors.

Activation loop

Protein kinases contain a large flexible loop, referred to as the activation loop (A-loop) whose conformation is postulated to regulate kinase activity (Figure 2). In many kinases the conformation of the AL is controlled by the phosphorylation of specific A-loop residues (Johnson et al., 1996). The loop can be generally defined as beginning with the conserved residues DFG and ending at the conserved APE sequence (Johnson et al., 1996). In VEGFR-2 this segment corresponds to D1046-E1075 and contains two tyrosines (Y1054 and Y1059). Both Y1054 and Y1059 were found to be autophosphorylation sites when the cytosolic domain of VEGFR-2 was expressed in *E. coli* (Dougher-Vermazen et al., 1994). Using the in vitro autophosphorylation protocol described here for VEGFR2Δ50, a stable phosphorylation site is indicated at Y1059, however no evidence of phosphorylation of Y1054 was detected.

In this unliganded VEGFR2Δ50P structure presented here, the A-loop appears quite mobile and interpretable electron density was not present for most of the central portion of the loop (G1048-G1063). This disorder is consistent with mobility of the A-loop deduced from other

kinase structures. For example, of the two molecules in the asymmetric unit of the unphosphorylated FGFR1 kinase structure the center of the A-loop has relatively high temperature factors in molecule A and is completely disordered in molecule B. Although residues 1048-1063 could not be modeled in VEGFR2Δ50P, unambiguous electron density was present for residues D1064-E1075, clearly indicating that these residues adopt a conformation similar to that observed in the unphosphorylated FGFR1 structure. The segment of D1064-P1068 has an extended structure that lies adjacent to the catalytic residues D1028 and R1032 (Figure 3a). Comparison to the structure of the (MgAMP-PNP)-peptide-IRKP complex structure indicates that the position of R1066-P1068 in this VEGFR2Δ50P structure is inhibitory to substrate binding. P1066 occupies equivalent space allocated to the tyrosine sidechain of the peptide substrate in the ternary IRK3P complex structure. The conformation of residues L1069-E1075 is similar to that in the ternary IRKP complex structure, however there is a complete directional change at P1068 (P1172 IRK) between the two structures. In the IRK structure residues N-terminal to this proline are directed toward αEF while in VEGFR2Δ50P they are directed toward αD on the opposite side of the protein (Figures 2 and 3).

Despite the phosphorylation of Y1059 prior to crystallization, the conformation seen here for residues D1064-P1068 is similar to the inhibitory conformation observed for analogous residues in the unphosphorylated FGFR1 structure. Y1059 in VEGFR2Δ50 corresponds to a relatively conserved phosphorylation site among protein tyrosine kinases. In the ternary IRKP complex structure and the phosphorylated lymphocyte kinase (Lck) structure (Yamaguchi and Hendrickson, 1996) the tyrosine at this position (Y1163 IRK, Y394 Lck) is phosphorylated and the A-loop has a non-inhibitory conformation similar to that observed in a phosphorylated cAPK ternary complex structure (Zheng et al., 1993). The interactions the phosphate group at this position makes with other protein residues are believed to help stabilize an A-loop conformation that allows substrate and ATP binding (Johnson et al., 1996; Hubbard, 1997). However, since this VEGFR2Δ50P structure described here does not exhibit a similar open A-loop conformation but rather has an inhibitory conformation with much of the loop disordered it is possible that the monophosphorylated A-loop of VEGFR2Δ50P exists in a dynamic equilibrium involving several conformations and that the conformation observed here is the one most favored in this crystal environment.

Kinase Insert Domain:

The kinase insert domain occurs in the kinase C-terminal lobe and connects helices αD and αE. In VEGFR-2 this region corresponds to a 68 residue stretch from N933 to L1000

(Figure 1). The lack of effect on intrinsic kinase activity (noted above) of deletion of residues T940-E989 is perhaps not surprising as the ends of the KID domain occur relatively far away (approximately 35-40 Å) from the catalytic site and on the opposite side of the protein from the position of the activation loop (Figure 2). These results are consistent with those for the CSF-1 receptor kinase in which deletion of 58 of the 64 residues of the CSF-1 KID only decreased it's ability to phosphorylate a peptide substrate by 10% (Taylor et al., 1989). Deletion of the entire 98 residues of βPDGFR, however, resulted in an 80% decrease in kinase activity towards a peptide substrate (Severinsson et al., (1990) Mol. Cell. Biol. *10*, 801-809). Thus, the present invention allows for the production of a synthetic catalytic linker which recognizes that the majority of KID is not required for catalysis but rather only a small number of residues must be present to form a linker between αD and αE so as to maintain a competent kinase structure.

In the VEGFR2Δ50P structure following αD, residues N933-P937 form a loose turn and an extended strand whose ends are roughly perpendicular to the axes of αD and αI at the C-terminus. In different Fourier maps, the electron density is strong and clear for residues N933-P937 and becomes weak for Y938 and K939 (sidechains of Y938 and K939 are not modeled) (Figure 5). The 50 residue deletion in VEGFR2Δ50 directly follows K939 so that the residue immediately C-terminal to K939 is V990, maintaining the residue numbering in full-length VEGFR-2. Residues V990-K997 are disordered and interpretable electron density begins again at D998. Residues D998-T1001 then form a short strand that joins αE at residue L1002 (Figures 5 and 6).

The two strands at the N-terminal and C-terminal ends of the KID form a pseudo two-stranded parallel β-sheet structure that is different from the conformations seen in this region of other kinase structures. The two ends of the KID thus make a variety of interactions which may help to stabilize the overall conformation and position of this domain in VEGFR-2. The sidechain of K931 makes an ionic interaction with the sidechain of E934 and also makes a hydrogen bond to the backbone carbonyl of D998 (Figure 6). Hydrogen bonding interactions between the strands include: E934 backbone C=O to L1000 NH, V936 NH to L1000 C=O, and P937 C=O to L1002 NH. In addition to these polar interactions, the sidechains of F935, P937, and L1000 are involved in extensive hydrophobic contacts. The sidechain of F935 is nestled in a hydrophobic pocket formed by the sidechains of L928, P937, L1000, L1002, L1005, L1101, and Y1130 (Figures 5 and 6). The L1000 sidechain also packs against the sidechains of Y927, K931, H1004, and Y1008.

It has been found by the applicants that deletion of portions of the KID also impart other useful and desirable characteristics to the modified VEGFR-2 polyprotein. The modified polypeptide has exhibited greater stability when exposed to higher temperatures in solution than the wild-type protein. Additionally, the modified polypeptide has also exhibited improved solubility than the wild-type protein. It is apparent to those skilled in the art that these properties allow improvements in various commercial aspects of the present invention. Examples of potential uses for the modified proteins include high-throughput screening of potential ligands for the receptor by various methods including those based on gene-chip technology (Affymax, Inc,) phage-display peptide libraries (The Ph.D. Kit® by New England BioLabs, Inc.) as well as in-depth analysis via FT-NMR.

It is therefore contemplated that the entire KID can be deleted and retain some catalytic activity in other related RTKs such as but not limited to PDGFα and β and other previously mentioned proteins. Furthermore, in one embodiment of the invention the entire KID is deleted and replaced with a synthetic catalytic linker of at least one amino acids such that both the catalytic activity and the crystallizability of the protein is retained.

Cloning of The PDGFRα Protein

In this example, the PDGFRα polyprotein is cloned using the methods outlined for VEGFR-2 above. The coding sequence for PDGFRα is derived from the sequence disclosed by Matsui,T., et al., (1989) Science 243: 800-804 (Accession No. 66814). PCR oligonucleotide primers are then made which code for residues located in the cytoplasmic domain and the catalytic domain of the protein. The catalytic domain of PDGFRα is shown starting at residue 689 (N) and ending at residue 791 (T) in Figure 1.

The remainder of the cloning and purification steps would be similar to those disclosed for the VEGFR2Δ50 protein and use technology well known to those skilled in the art.

It is contemplated that other members of the RTK family and other uses for the data disclosed herein and are not limited by the examples shown.